#9

INPUT SET: S31474.raw

> This Raw Listing contains the General
> Information Section and up to the first 5 pages.

SEQUENCE LISTING   **ENTERED**

```
1
2
3    (1)     General Information:
4
5         (i) APPLICANT: Malcolm, Bruce
6                        Taremi, Shahriar S.
7                        Weber, Patricia
8                        Yao, Nanhua
9
10      (ii) TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
11              NS3 Protease and NS4A Cofactor Peptide
12
13     (iii) NUMBER OF SEQUENCES: 123
14
15      (iv) CORRESPONDENCE ADDRESS:
16              (A) ADDRESSEE: Schering-Plough Corp.
17              (B) STREET: 2000 Galloping Hill Road
18              (C) CITY: Kenilworth
19              (D) STATE: New Jersey
20              (E) COUNTRY: USA
21              (F) ZIP: 07030
22
23       (v) COMPUTER READABLE FORM:
24              (A) MEDIUM TYPE: Floppy disk
25              (B) COMPUTER: Power Macintosh
26              (C) OPERATING SYSTEM: 8.0.1
27              (D) SOFTWARE: Microsoft Word 6.0.1
28
29      (vi) CURRENT APPLICATION DATA:
30              (A) APPLICATION NUMBER: US 09/198,723
31              (B) FILING DATE: 24 NOV 1998
32              (C) CLASSIFICATION:
33
34    (viii) ATTORNEY/AGENT INFORMATION:
35              (A) NAME: McLaughlin, Jaye P.
36              (B) REGISTRATION NUMBER: 41,211
37              (C) REFERENCE/DOCKET NUMBER: JB0800
38
39      (ix) TELECOMMUNICATION INFORMATION:
40              (A) TELEPHONE: (908)298-5056
41              (B) TELEFAX: (908)298-5388
42
43
44
45
46
```

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/09/198,723A*

INPUT SET: S31474.raw

```
47     (2) INFORMATION FOR SEQ ID NO:1:
48
49          (i) SEQUENCE CHARACTERISTICS:
50               (A) LENGTH: 216 amino acids
51               (B) TYPE: amino acid
52               (D) TOPOLOGY: linear
53
54          (ii) MOLECULE TYPE: protein
55
56          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
57
58     Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
59      1           5                  10                  15
60
61     Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
62              20          25                  30
63
64     Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu
65          35              40                  45
66
67     Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val
68          50              55                  60
69
70     Glu Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala
71      65              70                  75                  80
72
73     Thr Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser
74              85                  90                  95
75
76     Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn
77              100             105                 110
78
79     Val Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser
80          115                 120                 125
81
82     Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg
83          130             135                 140
84
85     His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser
86     145                 150                 155                 160
87
88     Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly
89                  165             170                 175
90
91     Pro Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala
92              180             185                 190
93
94     Val Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu
95              195             200                 205
96
97     Ser Met Glu Thr Thr Met Arg Ser
98          210             215
99
```

INPUT SET: S31474.raw

```
100
101
102
103        (2) INFORMATION FOR SEQ ID NO:2:
104
105            (i) SEQUENCE CHARACTERISTICS:
106                (A) LENGTH: 216 amino acids
107                (B) TYPE: amino acid
108                (D) TOPOLOGY: linear
109
110           (ii) MOLECULE TYPE: protein
111
112           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
113
114    Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
115     1               5                  10               15
116
117    Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
118             20                  25                  30
119
120    Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu
121             35                  40                  45
122
123    Leu Gly Cys Lys Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val
124         50                  55                  60
125
126    Glu Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala
127     65                  70                  75                  80
128
129    Thr Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser
130                 85                  90                  95
131
132    Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn
133                 100                 105                 110
134
135    Val Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser
136             115                 120                 125
137
138    Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg
139         130                 135                 140
140
141    His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser
142     145                 150                 155                 160
143
144    Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly
145                 165                 170                 175
146
147    Pro Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala
148                 180                 185                 190
149
150    Val Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu
151                 195                 200                 205
152
```

INPUT SET: S31474.raw

```
153    Ser Met Glu Thr Thr Met Arg Ser
154        210             215
155
156
157
158
159    (2) INFORMATION FOR SEQ ID NO:3:
160
161            (i) SEQUENCE CHARACTERISTICS:
162                    (A) LENGTH: 216 amino acids
163                    (B) TYPE: amino acid
164                    (D) TOPOLOGY: linear
165
166            (ii) MOLECULE TYPE: protein
167
168            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
169
170    Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
171     1           5               10              15
172
173    Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
174            20              25                  30
175
176    Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu
177            35              40                  45
178
179    Leu Gly Cys Ile Lys Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val
180        50              55                  60
181
182    Glu Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala
183     65              70                  75                  80
184
185    Thr Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser
186                85              90                  95
187
188    Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn
189                100             105                 110
190
191    Val Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser
192        115             120                 125
193
194    Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg
195        130             135                 140
196
197    His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser
198    145             150                 155                 160
199
200    Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly
201                165             170                 175
202
203    Pro Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala
204                180             185                 190
205
```

INPUT SET: S31474.raw

```
206   Val Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu
207       195                 200                 205

208
209   Ser Met Glu Thr Thr Met Arg Ser
210       210                 215

211
212
213
214
215   (2) INFORMATION FOR SEQ ID NO:4:
216
217         (i) SEQUENCE CHARACTERISTICS:
218               (A) LENGTH: 216 amino acids
219               (B) TYPE: amino acid
220               (D) TOPOLOGY: linear
221
222         (ii) MOLECULE TYPE: protein
223
224         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
225
226   Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
227    1               5                 10                 15

228
229   Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
230               20                 25                 30

231
232   Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu
233               35                 40                 45

234
235   Leu Gly Cys Lys Lys Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val
236           50                 55                 60

237
238   Glu Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala
239    65                 70                 75                 80

240
241   Thr Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser
242               85                 90                 95

243
244   Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn
245               100                 105                 110

246
247   Val Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser
248           115                 120                 125

249
250   Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg
251           130                 135                 140

252
253   His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser
254    145                 150                 155                 160

255
256   Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly
257               165                 170                 175

258
```

***** PREVIOUSLY ERRORED SEQUENCES - EDITED *****

```
4116    (2) INFORMATION FOR SEQ ID NO:92:
4117
4118        (i) SEQUENCE CHARACTERISTICS:
4119            (A) LENGTH: 651 base pairs
4120            (B) TYPE: nucleic acid
4121            (C) STRANDEDNESS: single
4122            (D) TOPOLOGY: linear
4123
4124    (ii) MOLECULE TYPE: cDNA
4125
4126
4127
4128    (ix) FEATURE:
4129            (A) NAME/KEY: CDS
4130            (B) LOCATION: 1..651
4131
4132
4133    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:
4134
4135    ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG      48
4136    Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
4137     1               5                   10                  15
4138
4139    CGC GGC AGC CAT ATG GGT TCT GTT GTT ATT GTT GGT AGA ATT ATT TTA      96
4140    Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
4141             20                  25                  30
4142
4143    TCT GGT AGT GGT AGT ATC ACG GCC TAC TCC CAA CAG ACG CGG GGC CTA     144
4144    Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu
4145             35                  40                  45
4146
4147    CTT GGT TGC ATC ATC ACT AGC CTT ACA GGC CGG GAC AAG AAC CAG GTC     192
4148    Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val
4149         50                  55                  60
4150
4151    GAG GGA GAG GTT CAG GTG GTT TCC ACC GCA ACA CAA TCC TTC CTG GCG     240
4152    Glu Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala
4153     65                  70                  75                  80
4154
4155    ACC TGC GTC AAC GGC GTG TGT TGG ACC GTT TAC CAT GGT GCT GGC TCA     288
4156    Thr Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser
4157                     85                  90                  95
4158
4159    AAG ACC TTA GCC GGC CCA AAG GGG CCA ATC ACC CAG ATG TAC ACT AAT     336
4160    Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn
4161                 100                 105                 110
4162
```

INPUT SET: S31474.raw

```
4163    GTG GAC CAG GAC CTC GTC GGC TGG CAG GCG CCC CCC GGG GCG CGT TCC        384
4164    Val Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser
4165        115                     120                 125
4166
4167    TTG ACA CCA TGC ACC TGT GGC AGC TCA GAC CTT TAC TTG GTC ACG AGA        432
4168    Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg
4169        130                     135                 140
4170
4171    CAT GCT GAC GTC ATT CCG GTG CGC CGG CGG GGC GAC AGT AGG GGG AGC        480
4172    His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser
4173    145                 150                 155                 160
4174
4175    CTG CTC TCC CCC AGG CCT GTC TCC TAC TTG AAG GGC TCT TCG GGT GGT        528
4176    Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly
4177                165                 170                 175
4178
4179    CCA CTG CTC TGC CCT TCG GGG CAC GCT GTG GGC ATC TTC CGG GCT GCC        576
4180    Pro Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala
4181        .       180                 185                 190
4182
4183    GTA TGC ACC CGG GGG GTT GCG AAG GCG GTG GAC TTT GTG CCC GTA GAG        624
4184    Val Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu
4185        195                 200                 205
4186
4187    TCC ATG GAA ACT ACT ATG CGG TCT TGA                                    651
4188    Ser Met Glu Thr Thr Met Arg Ser
4189        210                 215
4190
4191
```

```
4869    (2) INFORMATION FOR SEQ ID NO:102:
4870
4871        (i) SEQUENCE CHARACTERISTICS:
4872            (A) LENGTH: 1998 base pairs
4873            (B) TYPE: nucleic acid
4874            (C) STRANDEDNESS: single
4875            (D) TOPOLOGY: linear
4876
4877       (ii) MOLECULE TYPE: cDNA
4878
4879
4880       (ix) FEATURE:
4881            (A) NAME/KEY: CDS
4882            (B) LOCATION: 1..1998
4883
4884
4885       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:
4886
4887    ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG        48
4888    Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
4889    1                   5                   10                  15
4890
4891    CGC GGC AGC CAT ATG GGT TCT GTT GTT ATT GTT GGT AGA ATT ATT TTA        96
```

INPUT SET: S31474.raw

```
4892    Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
4893             20          25              30
4894
4895    TCT GGT AGT GGT AGT ATC ACG GCC TAC TCC CAA CAG ACG CGG GGC CTA    144
4896    Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu
4897             35          40              45
4898
4899    CTT GGT TGC ATC ATC ACT AGC CTT ACA GGC CGG GAC AAG AAC CAG GTC    192
4900    Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val
4901             50          55              60
4902
4903    GAG GGA GAG GTT CAG GTG GTT TCC ACC GCA ACA CAA TCC TTC CTG GCG    240
4904    Glu Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala
4905    65              70              75              80
4906
4907    ACC TGC GTC AAC GGC GTG TGT TGG ACC GTT TAC CAT GGT GCT GGC TCA    288
4908    Thr Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser
4909             85              90              95
4910
4911    AAG ACC TTA GCC GGC CCA AAG GGG CCA ATC ACC CAG ATG TAC ACT AAT    336
4912    Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn
4913             100             105             110
4914
4915    GTG GAC CAG GAC CTC GTC GGC TGG CAG GCG CCC CCC GGG GCG CGT TCC    384
4916    Val Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser
4917             115             120             125
4918
4919    TTG ACA CCA TGC ACC TGT GGC AGC TCA GAC CTT TAC TTG GTC ACG AGA    432
4920    Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg
4921             130             135             140
4922
4923    CAT GCT GAC GTC ATT CCG GTG CGC CGG CGG GGC GAC AGT AGG GGG AGC    480
4924    His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser
4925    145             150             155             160
4926
4927    CTG CTC TCC CCC AGG CCT GTC TCC TAC TTG AAG GGC TCT TCG GGT GGT    528
4928    Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly
4929             165             170             175
4930
4931    CCA CTG CTC TGC CCT TCG GGG CAC GCT GTG GGC ATC TTC CGG GCT GCC    576
4932    Pro Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala
4933             180             185             190
4934
4935    GTA TGC ACC CGG GGG GTT GCG AAG GCG GTG GAC TTT GTG CCC GTA GAG    624
4936    Val Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu
4937             195             200             205
4938
4939    TCC ATG GAA ACT ACT ATG CGG TCT CCG GTC TTC ACG GAC AAC TCA TCC    672
4940    Ser Met Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser
4941    210             215             220
4942
4943    CCC CCG GCC GTA CCG CAG TCA TTT CAA GTG GCC CAC CTA CAC GCT CCC    720
4944    Pro Pro Ala Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/198,723A*

INPUT SET: S31474.raw

```
4945   225                    230                    235                    240
4946
4947   ACT GGC AGC GGC AAG AGT ACT AAA GTG CCG GCT GCA TAT GCA GCC CAA        768
4948   Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln
4949                   245                    250                    255
4950
4951   GGG TAC AAG GTG CTC GTC CTC AAT CCG TCC GTT GCC GCT ACC TTA GGG        816
4952   Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly
4953                   260                    265                    270
4954
4955   TTT GGG GCG TAT ATG TCT AAG GCA CAC GGT ATT GAC CCC AAC ATC AGA        864
4956   Phe Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg
4957                   275                    280                    285
4958
4959   ACT GGG GTA AGG ACC ATT ACC ACA GGC GCC CCC GTC ACA TAC TCT ACC        912
4960   Thr Gly Val Arg Thr Ile Thr Thr Gly Ala Pro Val Thr Tyr Ser Thr
4961          290                    295                    300
4962
4963   TAT GGC AAG TTT CTT GCC GAT GGT GGT TGC TCT GGG GGC GCT TAT GAC        960
4964   Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp
4965   305                    310                    315                    320
4966
4967   ATC ATA ATA TGT GAT GAG TGC CAT TCA ACT GAC TCG ACT ACA ATC TTG       1008
4968   Ile Ile Ile Cys Asp Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu
4969                   325                    330                    335
4970
4971   GGC ATC GGC ACA GTC CTG GAC CAA GCG GAG ACG GCT GGA GCG CGG CTT       1056
4972   Gly Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu
4973                   340                    345                    350
4974
4975   GTC GTG CTC GCC ACC GCT ACG CCT CCG GGA TCG GTC ACC GTG CCA CAC       1104
4976   Val Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His
4977          355                    360                    365
4978
4979   CCA AAC ATC GAG GAG GTG GCC CTG TCT AAT ACT GGA GAG ATC CCC TTC       1152
4980   Pro Asn Ile Glu Glu Val Ala Leu Ser Asn Thr Gly Glu Ile Pro Phe
4981          370                    375                    380
4982
4983   TAT GGC AAA GCC ATC CCC ATT GAA GCC ATC AGG GGG GGA AGG CAT CTC       1200
4984   Tyr Gly Lys Ala Ile Pro Ile Glu Ala Ile Arg Gly Gly Arg His Leu
4985   385                    390                    395                    400
4986
4987   ATT TTC TGT CAT TCC AAG AAG AAG TGC GAC GAG CTC GCC GCA AAG CTG       1248
4988   Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu
4989                   405                    410                    415
4990
4991   TCA GGC CTC GGA ATC AAC GCT GTG GCG TAT TAC CGG GGG CTC GAT GTG       1296
4992   Ser Gly Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val
4993                   420                    425                    430
4994
4995   TCC GTC ATA CCA ACT ATC GGA GAC GTC GTT GTC GTG GCA ACA GAC GCT       1344
4996   Ser Val Ile Pro Thr Ile Gly Asp Val Val Val Val Ala Thr Asp Ala
4997          435                    440                    445
```

APR 1 7 2000

INPUT SET: S31474.raw

```
4998
4999   CTG ATG ACG GGC TAT ACG GGC GAC TTT GAC TCA GTG ATC GAC TGT AAC      1392
5000   Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn
5001       450             455             460
5002
5003   ACA TGT GTC ACC CAG ACA GTC GAC TTC AGC TTG GAT CCC ACC TTC ACC      1440
5004   Thr Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr
5005   465             470             475             480
5006
5007   ATT GAG ACG ACG ACC GTG CCT CAA GAC GCA GTG TCG CGC TCG CAG CGG      1488
5008   Ile Glu Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg
5009               485             490             495
5010
5011   CGG GGT AGG ACT GGC AGG GGT AGG AGA GGC ATC TAC AGG TTT GTG ACT      1536
5012   Arg Gly Arg Thr Gly Arg Gly Arg Arg Gly Ile Tyr Arg Phe Val Thr
5013               500             505             510
5014
5015   CCG GGA GAA CGG CCC TCG GGC ATG TTC GAT TCC TCG GTC CTG TGT GAG      1584
5016   Pro Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu
5017           515             520             525
5018
5019   TGC TAT GAC GCG GGC TGT GCT TGG TAC GAG CTC ACC CCC GCC GAG ACC      1632
5020   Cys Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr
5021       530             535             540
5022
5023   TCG GTT AGG TTG CGG GCC TAC CTG AAC ACA CCA GGG TTG CCC GTT TGC      1680
5024   Ser Val Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys
5025   545             550             555             560
5026
5027   CAG GAC CAC CTG GAG TTC TGG GAG AGT GTC TTC ACA GGC CTC ACC CAT      1728
5028   Gln Asp His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His
5029               565             570             575
5030
5031   ATA GAT GCA CAC TTC TTG TCC CAG ACC AAG CAG GCA GGA GAC AAC TTC      1776
5032   Ile Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe
5033           580             585             590
5034
5035   CCC TAC CTG GTA GCA TAC CAA GCC ACG GTG TGC GCC AGG GCT CAG GCC      1824
5036   Pro Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala
5037           595             600             605
5038
5039   CCA CCT CCA TCA TGG GAT CAA ATG TGG AAG TGT CTC ATA CGG CTG AAA      1872
5040   Pro Pro Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys
5041       610             615             620
5042
5043   CCT ACG CTG CAC GGG CCA ACA CCC TTG CTG TAC AGG CTG GGA GCC GTC      1920
5044   Pro Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val
5045   625             630             635             640
5046
5047   CAA AAT GAG GTC ACC CTC ACC CAC CCC ATA ACC AAA TAC ATC ATG GCA      1968
5048   Gln Asn Glu Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala
5049               645             650             655
5050
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/198,723A*

INPUT SET: S31474.raw

```
5051   TGC ATG TCG GCT GAC CTG GAG GTC GTC ACT                        1998
5052   Cys Met Ser Ala Asp Leu Glu Val Val Thr
5053           660                 665
5054
5055
5056
```

```
6561   (2) INFORMATION FOR SEQ ID NO:111:
6562
6563        (i) SEQUENCE CHARACTERISTICS:
6564             (A) LENGTH: 2016 base pairs
6565             (B) TYPE: nucleic acid
6566             (C) STRANDEDNESS: single
6567             (D) TOPOLOGY: linear
6568
6569        (ii) MOLECULE TYPE: cDNA
6570
6571
6572        (ix) FEATURE:
6573             (A) NAME/KEY: CDS
6574             (B) LOCATION: 1..2016
6575
6576
6577        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:
6578
6579   ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG        48
6580   Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
6581   1               5                   10                  15
6582
6583   CGC GGC AGC CAT ATG GCT TAC TCT CTG ACT ACG GGT TCT GTT GTT ATT        96
6584   Arg Gly Ser His Met Ala Tyr Ser Leu Thr Thr Gly Ser Val Val Ile
6585               20                  25                  30
6586
6587   GTT GGT AGA ATT ATT TTA TCT GGT AGT GGT AGT ATC ACG GCC TAC TCC       144
6588   Val Gly Arg Ile Ile Leu Ser Gly Ser Gly Ser Ile Thr Ala Tyr Ser
6589           35                  40                  45
6590
6591   CAA CAG ACG CGG GGC CTA CTT GGT TGC ATC ATC ACT AGC CTT ACA GGC       192
6592   Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly
6593           50                  55                  60
6594
6595   CGG GAC AAG AAC CAG GTC GAG GGA GAG GTT CAG GTG GTT TCC ACC GCA       240
6596   Arg Asp Lys Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr Ala
6597   65                  70                  75                  80
6598
6599   ACA CAA TCC TTC CTG GCG ACC TGC GTC AAC GGC GTG TGT TGG ACC GTT       288
6600   Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr Val
6601               85                  90                  95
6602
6603   TAC CAT GGT GCT GGC TCA AAG ACC TTA GCC GGC CCA AAG GGG CCA ATC       336
6604   Tyr His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile
6605               100                 105                 110
6606
```

INPUT SET: S31474.raw

```
6607   ACC CAG ATG TAC ACT AAT GTG GAC CAG GAC CTC GTC GGC TGG CAG GCG          384
6608   Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln Ala
6609           115                 120                 125
6610
6611   CCC CCC GGG GCG CGT TCC TTG ACA CCA TGC ACC TGT GGC AGC TCA GAC          432
6612   Pro Pro Gly Ala Arg Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp
6613           130                 135                 140
6614
6615   CTT TAC TTG GTC ACG AGA CAT GCT GAC GTC ATT CCG GTG CGC CGG CGG          480
6616   Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg Arg
6617   145                 150                 155                 160
6618
6619   GGC GAC AGT AGG GGG AGC CTG CTC TCC CCC AGG CCT GTC TCC TAC TTG          528
6620   Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu
6621               165                 170                 175
6622
6623   AAG GGC TCT GCT GGT GGT CCA CTG CTC TGC CCT TCG GGG CAC GCT GTG          576
6624   Lys Gly Ser Ala Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala Val
6625               180                 185                 190
6626
6627   GGC ATC TTC CGG GCT GCC GTA TGC ACC CGG GGG GTT GCG AAG GCG GTG          624
6628   Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala Val
6629           195                 200                 205
6630
6631   GAC TTT GTG CCC GTA GAG TCC ATG GAA ACT ACT ATG CGG TCT CCG GTC          672
6632   Asp Phe Val Pro Val Glu Ser Met Glu Thr Thr Met Arg Ser Pro Val
6633           210                 215                 220
6634
6635   TTC ACG GAC AAC TCA TCC CCC CCG GCC GTA CCG CAG TCA TTT CAA GTG          720
6636   Phe Thr Asp Asn Ser Ser Pro Pro Ala Val Pro Gln Ser Phe Gln Val
6637   225                 230                 235                 240
6638
6639   GCC CAC CTA CAC GCT CCC ACT GGC AGC GGC AAG AGT ACT AAA GTG CCG          768
6640   Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro
6641               245                 250                 255
6642
6643   GCT GCA TAT GCA GCC CAA GGG TAC AAG GTG CTC GTC CTC AAT CCG TCC          816
6644   Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser
6645               260                 265                 270
6646
6647   GTT GCC GCT ACC TTA GGG TTT GGG GCG TAT ATG TCT AAG GCA CAC GGT          864
6648   Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly
6649           275                 280                 285
6650
6651   ATT GAC CCC AAC ATC AGA ACT GGG GTA AGG ACC ATT ACC ACA GGC GCC          912
6652   Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ala
6653           290                 295                 300
6654
6655   CCC GTC ACA TAC TCT ACC TAT GGC AAG TTT CTT GCC GAT GGT GGT TGC          960
6656   Pro Val Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys
6657   305                 310                 315                 320
6658
6659   TCT GGG GGC GCT TAT GAC ATC ATA ATA TGT GAT GAG TGC CAT TCA ACT          1008
```

INPUT SET: S31474.raw

```
6660    Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr
6661                    325             330             335

6662
6663    GAC TCG ACT ACA ATC TTG GGC ATC GGC ACA GTC CTG GAC CAA GCG GAG    1056
6664    Asp Ser Thr Thr Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu
6665                340             345             350

6666
6667    ACG GCT GGA GCG CGG CTT GTC GTG CTC GCC ACC GCT ACG CCT CCG GGA    1104
6668    Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly
6669                355             360             365

6670
6671    TCG GTC ACC GTG CCA CAC CCA AAC ATC GAG GAG GTG GCC CTG TCT AAT    1152
6672    Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Asn
6673        370             375             380

6674
6675    ACT GGA GAG ATC CCC TTC TAT GGC AAA GCC ATC CCC ATT GAA GCC ATC    1200
6676    Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Ile Glu Ala Ile
6677    385                 390             395             400

6678
6679    AGG GGG GGA AGG CAT CTC ATT TTC TGT CAT TCC AAG AAG AAG TGC GAC    1248
6680    Arg Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp
6681                405             410             415

6682
6683    GAG CTC GCC GCA AAG CTG TCA GGC CTC GGA ATC AAC GCT GTG GCG TAT    1296
6684    Glu Leu Ala Ala Lys Leu Ser Gly Leu Gly Ile Asn Ala Val Ala Tyr
6685                420             425             430

6686
6687    TAC CGG GGG CTC GAT GTG TCC GTC ATA CCA ACT ATC GGA GAC GTC GTT    1344
6688    Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ile Gly Asp Val Val
6689                435             440             445

6690
6691    GTC GTG GCA ACA GAC GCT CTG ATG ACG GGC TAT ACG GGC GAC TTT GAC    1392
6692    Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp
6693        450             455             460

6694
6695    TCA GTG ATC GAC TGT AAC ACA TGT GTC ACC CAG ACA GTC GAC TTC AGC    1440
6696    Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser
6697    465             470             475             480

6698
6699    TTG GAT CCC ACC TTC ACC ATT GAG ACG ACG ACC GTG CCT CAA GAC GCA    1488
6700    Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Val Pro Gln Asp Ala
6701                485             490             495

6702
6703    GTG TCG CGC TCG CAG CGG CGG GGT AGG ACT GGC AGG GGT AGG AGA GGC    1536
6704    Val Ser Arg Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Arg Gly
6705                500             505             510

6706
6707    ATC TAC AGG TTT GTG ACT CCG GGA GAA CGG CCC TCG GGC ATG TTC GAT    1584
6708    Ile Tyr Arg Phe Val Thr Pro Gly Glu Arg Pro Ser Gly Met Phe Asp
6709                515             520             525

6710
6711    TCC TCG GTC CTG TGT GAG TGC TAT GAC GCG GGC TGT GCT TGG TAC GAG    1632
6712    Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr Glu
```

INPUT SET: S31474.raw

```
6713            530                      535                      540
6714
6715   CTC ACC CCC GCC GAG ACC TCG GTT AGG TTG CGG GCC TAC CTG AAC ACA      1680
6716   Leu Thr Pro Ala Glu Thr Ser Val Arg Leu Arg Ala Tyr Leu Asn Thr
6717   545                      550                      555                      560
6718
6719   CCA GGG TTG CCC GTT TGC CAG GAC CAC CTG GAG TTC TGG GAG AGT GTC      1728
6720   Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ser Val
6721                   565                      570                      575
6722
6723   TTC ACA GGC CTC ACC CAT ATA GAT GCA CAC TTC TTG TCC CAG ACC AAG      1776
6724   Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr Lys
6725                   580                      585                      590
6726
6727   CAG GCA GGA GAC AAC TTC CCC TAC CTG GTA GCA TAC CAA GCC ACG GTG      1824
6728   Gln Ala Gly Asp Asn Phe Pro Tyr Leu Val Ala Tyr Gln Ala Thr Val
6729                   595                      600                      605
6730
6731   TGC GCC AGG GCT CAG GCC CCA CCT CCA TCA TGG GAT CAA ATG TGG AAG      1872
6732   Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp Lys
6733                   610                      615                      620
6734
6735   TGT CTC ATA CGG CTG AAA CCT ACG CTG CAC GGG CCA ACA CCC TTG CTG      1920
6736   Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu Leu
6737   625                      630                      635                      640
6738
6739   TAC AGG CTG GGA GCC GTC CAA AAT GAG GTC ACC CTC ACC CAC CCC ATA      1968
6740   Tyr Arg Leu Gly Ala Val Gln Asn Glu Val Thr Leu Thr His Pro Ile
6741                   645                      650                      655
6742
6743   ACC AAA TAC ATC ATG GCA TGC ATG TCG GCT GAC CTG GAG GTC GTC ACT      2016
6744   Thr Lys Tyr Ile Met Ala Cys Met Ser Ala Asp Leu Glu Val Val Thr
6745                   660                      665                      670
6746
6747
6748
6749
6750
```

```
6751   (2) INFORMATION FOR SEQ ID NO:112:
6752
6753        (i) SEQUENCE CHARACTERISTICS:
6754            (A) LENGTH: 648 base pairs
6755            (B) TYPE: nucleic acid
6756            (C) STRANDEDNESS: single
6757            (D) TOPOLOGY: linear
6758
6759       (ii) MOLECULE TYPE: cDNA
6760
6761
6762       (ix) FEATURE:
6763            (A) NAME/KEY: CDS
6764            (B) LOCATION: 1..648
```

```
6765
6766
6767        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:
6768
6769   ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG        48
6770   Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
6771    1               5                   10                  15
6772
6773   CGC GGC AGC CAT ATG GGT TCT GTT GTT ATT GTT GGT AGA ATT ATT TTA        96
6774   Arg Gly Ser His Met Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu
6775                   20                  25                  30
6776
6777   TCT CCT GCT GGT ATC ACG GCC TAC TCC CAA CAG ACG CGG GGC CTA CTT       144
6778   Ser Pro Ala Gly Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu Leu
6779               35                  40                  45
6780
6781   GGT TGC ATC ATC ACT AGC CTT ACA GGC CGG GAC AAG AAC CAG GTC GAG       192
6782   Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu
6783           50                  55                  60
6784
6785   GGA GAG GTT CAG GTG GTT TCC ACC GCA ACA CAA TCC TTC CTG GCG ACC       240
6786   Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr
6787    65                  70                  75                  80
6788
6789   TGC GTC AAC GGC GTG TGT TGG ACC GTT TAC CAT GGT GCT GGC TCA AAG       288
6790   Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys
6791                   85                  90                  95
6792
```

```
6793   ACC TTA GCC GGC CCA AAG GGG CCA ATC ACC CAG ATG TAC ACT AAT GTG        336
6794   Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val
6795               100                 105                 110
6796
6797   GAC CAG GAC CTC GTC GGC TGG CAG GCG CCC CCC GGG GCG CGT TCC TTG        384
6798   Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Leu
6799               115                 120                 125
6800
6801   ACA CCA TGC ACC TGT GGC AGC TCA GAC CTT TAC TTG GTC ACG AGA CAT        432
6802   Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His
6803           130                 135                 140
6804
6805   GCT GAC GTC ATT CCG GTG CGC CGG CGG GGC GAC AGT AGG GGG AGC CTG        480
6806   Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu
6807   145                 150                 155                 160
6808
6809   CTC TCC CCC AGG CCT GTC TCC TAC TTG AAG GGC TCT TCG GGT GGT CCA        528
6810   Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro
6811               165                 170                 175
6812
6813   CTG CTC TGC CCT TCG GGG CAC GCT GTG GGC ATC TTC CGG GCT GCC GTA        576
6814   Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala Val
6815           180                 185                 190
6816
6817   TGC ACC CGG GGG GTT GCG AAG GCG GTG GAC TTT GTG CCC GTA GAG TCC        624
```

INPUT SET: S31474.raw

```
6818   Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu Ser
6819           195             200             205
6820
6821   ATG GAA ACT ACT ATG CGG TCT TGA                                    648
6822   Met Glu Thr Thr Met Arg Ser
6823       210             215
6824    ·
6825
6826
```

```
7236   (2) INFORMATION FOR SEQ ID NO:117:
7237
7238       (i) SEQUENCE CHARACTERISTICS:
7239           (A) LENGTH: 1941 base pairs
7240           (B) TYPE: nucleic acid
7241           (C) STRANDEDNESS: single
7242           (D) TOPOLOGY: linear
7243
7244       (ii) MOLECULE TYPE: cDNA
7245
7246
7247
7248       (ix) FEATURE:
7249           (A) NAME/KEY: CDS
7250           (B) LOCATION: 1..1941
7251
7252
7253       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:
7254
7255   ATG CAT ATG CAT CAT CAT CAC CAT CAT CTG GTG CCG CGC GGC AGC GCG      48
7256   Met His Met His His His His His His Leu Val Pro Arg Gly Ser Ala
7257    1               5                   10                  15
7258
7259   CCC ATC ACG GCC TAC TCC CAA CAG ACG CGG GGC CTA CTT GGT TGC ATC      96
7260   Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile
7261               20                  25                  30
7262
7263   AAG ACT AGC CTT ACA GGC CGG GAC AAG AAC CAG GTC GAG GGA GAG GTT      144
7264   Lys Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu Val
7265           35                  40              .45
7266
7267   CAG GTG GTT TCC ACC GCA ACA CAA TCC TTC CTG GCG ACC TGC GTC AAC      192
7268   Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr Cys Val Asn
7269       50                  55                  60
7270
7271   GGC GTG TGT TGG ACC GTT TAC CAT GGT GCT GGC TCA AAG ACC TTA GCC      240
7272   Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys Thr Leu Ala
7273    65 ·                70                  75                  80
7274
7275   GGC CCA AAG GGG CCA ATC ACC CAG ATG TAC ACT AAT GTG GAC CAG GAC      288
7276   Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val Asp Gln Asp
7277               85                  90                  95
7278
```

INPUT SET: S31474.raw

```
7279  CTC GTC GGC TGG CAG GCG CCC CCC GGG GCG CGT TCC TTG ACA CCA TGC        336
7280  Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Leu Thr Pro Cys
7281          100                     105                 110
7282
7283  ACC TGT GGC AGC TCA GAC CTT TAC TTG GTC ACG AGA CAT GCT GAC GTC        384
7284  Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val
7285          115                     120                 125
7286
7287  ATT CCG GTG CGC CGG CGG GGC GAC AGT AGG GGG AGC CTG CTC TCC CCC        432
7288  Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro
7289          130                     135                 140
7290
7291  AGG CCT GTC TCC TAC TTG AAG GGC TCT TCG GGT GGT CCA CTG CTC TGC        480
7292  Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys
7293  145                     150                 155                 160
7294
7295  CCT TCG GGG CAC GCT GTG GGC ATC TTC CGG GCT GCC GTA TGC ACC CGG        528
7296  Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys Thr Arg
7297                  165                 170                 175
7298
7299  GGG GTT GCG AAG GCG GTG GAC TTT GTG CCC GTA GAG TCC ATG GAA ACT        576
7300  Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu Ser Met Glu Thr
7301              180                     185                 190
7302
7303  ACT ATG CGG TCT CCG GTC TTC ACG GAC AAC TCA TCC CCC CCG GCC GTA        624
7304  Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Ala Val
7305              195                     200                 205
7306
7307  CCG CAG TCA TTT CAA GTG GCC CAC CTA CAC GCT CCC ACT GGC AGC GGC        672
7308  Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly Ser Gly
7309          210                     215                 220
7310
7311  AAG AGT ACT AAA GTG CCG GCT GCA TAT GCA GCC CAA GGG TAC AAG GTG        720
7312  Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val
7313  225                     230                 235                 240
7314
7315  CTC GTC CTC AAT CCG TCC GTT GCC GCT ACC TTA GGG TTT GGG GCG TAT        768
7316  Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr
7317                  245                 250                 255
7318
7319  ATG TCT AAG GCA CAC GGT ATT GAC CCC AAC ATC AGA ACT GGG GTA AGG        816
7320  Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg
7321              260                     265                 270
7322
7323  ACC ATT ACC ACA GGC GCC CCC GTC ACA TAC TCT ACC TAT GGC AAG TTT        864
7324  Thr Ile Thr Thr Gly Ala Pro Val Thr Tyr Ser Thr Tyr Gly Lys Phe
7325              275                     280                 285
7326
7327  CTT GCC GAT GGT GGT TGC TCT GGG GGC GCT TAT GAC ATC ATA ATA TGT        912
7328  Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys
7329          290                     295                 300
7330
7331  GAT GAG TGC CAT TCA ACT GAC TCG ACT ACA ATC TTG GGC ATC GGC ACA        960
```

INPUT SET: S31474.raw

```
7332   Asp Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu Gly Ile Gly Thr
7333   305                 310             315                 320
7334
7335   GTC CTG GAC CAA GCG GAG ACG GCT GGA GCG CGG CTT GTC GTG CTC GCC    1008
7336   Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala
7337                   325             330                 335
7338
7339   ACC GCT ACG CCT CCG GGA TCG GTC ACC GTG CCA CAC CCA AAC ATC GAG    1056
7340   Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu
7341                   340             345                 350
7342
7343   GAG GTG GCC CTG TCT AAT ACT GGA GAG ATC CCC TTC TAT GGC AAA GCC    1104
7344   Glu Val Ala Leu Ser Asn Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala
7345                   355             360                 365
7346
7347   ATC CCC ATT GAA GCC ATC AGG GGG GGA AGG CAT CTC ATT TTC TGT CAT    1152
7348   Ile Pro Ile Glu Ala Ile Arg Gly Gly Arg His Leu Ile Phe Cys His
7349       370             375             380
7350
7351   TCC AAG AAG AAG TGC GAC GAG CTC GCC GCA AAG CTG TCA GGC CTC GGA    1200
7352   Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Ser Gly Leu Gly
7353   385             390             395                 400
7354
7355   ATC AAC GCT GTG GCG TAT TAC CGG GGG CTC GAT GTG TCC GTC ATA CCA    1248
7356   Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro
7357                   405             410                 415
7358
7359   ACT ATC GGA GAC GTC GTT GTC GTG GCA ACA GAC GCT CTG ATG ACG GGC    1296
7360   Thr Ile Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met Thr Gly
7361               420             425                 430
7362
7363   TAT ACG GGC GAC TTT GAC TCA GTG ATC GAC TGT AAC ACA TGT GTC ACC    1344
7364   Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr
7365           435             440             445
7366
7367   CAG ACA GTC GAC TTC AGC TTG GAT CCC ACC TTC ACC ATT GAG ACG ACG    1392
7368   Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr
7369       450             455             460
7370
7371   ACC GTG CCT CAA GAC GCA GTG TCG CGC TCG CAG CGG CGG GGT AGG ACT    1440
7372   Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly Arg Thr
7373   465             470             475                 480
7374
7375   GGC AGG GGT AGG AGA GGC ATC TAC AGG TTT GTG ACT CCG GGA GAA CGG    1488
7376   Gly Arg Gly Arg Arg Gly Ile Tyr Arg Phe Val Thr Pro Gly Glu Arg
7377               485             490                 495
7378
7379   CCC TCG GGC ATG TTC GAT TCC TCG GTC CTG TGT GAG TGC TAT GAC GCG    1536
7380   Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala
7381           500             505             510
7382
7383   GGC TGT GCT TGG TAC GAG CTC ACC CCC GCC GAG ACC TCG GTT AGG TTG    1584
7384   Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser Val Arg Leu
```

INPUT SET: S31474.raw

```
7385           515                520                525
7386
7387   CGG GCC TAC CTG AAC ACA CCA GGG TTG CCC GTT TGC CAG GAC CAC CTG      1632
7388   Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu
7389        530                535                540
7390
7391   GAG TTC TGG GAG AGT GTC TTC ACA GGC CTC ACC CAT ATA GAT GCA CAC      1680
7392   Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile Asp Ala His
7393   545                550                555                560
7394
7395   TTC TTG TCC CAG ACC AAG CAG GCA GGA GAC AAC TTC CCC TAC CTG GTA      1728
7396   Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro Tyr Leu Val
7397            565                570                575
7398
7399   GCA TAC CAA GCC ACG GTG TGC GCC AGG GCT CAG GCC CCA CCT CCA TCA      1776
7400   Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser
7401            580                585                590
7402
7403   TGG GAT CAA ATG TGG AAG TGT CTC ATA CGG CTG AAA CCT ACG CTG CAC      1824
7404   Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His
7405            595                600                605
7406
7407   GGG CCA ACA CCC TTG CTG TAC AGG CTG GGA GCC GTC CAA AAT GAG GTC      1872
7408   Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Val
7409        610                615                620
7410
7411   ACC CTC ACC CAC CCC ATA ACC AAA TAC ATC ATG GCA TGC ATG TCG GCC      1920
7412   Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys Met Ser Ala
7413   625                630                635                640
7414
7415   GAC CTG GAG GTC GTT ACG TAG                                          1941
7416   Asp Leu Glu Val Val Thr
7417            645
7418
7419
7420
```

```
7974   (2) INFORMATION FOR SEQ ID NO:121:
7975
7976       (i) SEQUENCE CHARACTERISTICS:
7977           (A) LENGTH: 36 base pairs
7978           (B) TYPE: nucleic acid
7979           (C) STRANDEDNESS: single
7980           (D) TOPOLOGY: linear
7981
7982       (ii) MOLECULE TYPE: RNA (genomic)
7983
7984
7985
7986
7987       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:
7988
7989   GCUCGCCCGG GGAUCCUCUA GGAAUACACG UUCGAU            36
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/09/198,723A*

INPUT SET: S31474.raw

Line   Error   Original Text